

- 24 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANTS: Vinik, Aaron I.
Pittenger, Gary L.
Rafaeloff, Ronit
10 Rosenberg, Lawrence
Duguid, William P.
- (ii) TITLE OF INVENTION: INGAP PROTEIN INVOLVED IN PANCREATIC
15 ISLET NEOGENESIS
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
20 (A) ADDRESSEE: Banner & Witcoff, Ltd.
(B) STREET: 1001 G Street, N.W.
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: US
25 (F) ZIP: 20001-4597
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
30 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
35 (B) FILING DATE: 09-SEP-1996
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Kagan, Sarah A.
40 (B) REGISTRATION NUMBER: 32,141
(C) REFERENCE/DOCKET NUMBER: 00570.59178
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 202-508-9100
45 (B) TELEFAX: 202-508-9299

(2) INFORMATION FOR SEQ ID NO:1:

- 50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 747 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- 55 (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 60 (vi) ORIGINAL SOURCE:

- 25 -

(A) ORGANISM: *Cricetulus*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 20..541

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

10	CTGCAAGACA GGTACCATG ATG CTT CCC ATG ACC CTC TGT AGG ATG TCT TGG	52
	Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp	
	1 5 10	
15	ATG CTG CTT TCC TGC CTG ATG TTC CTT TCT TGG GTG GAA GGT GAA GAA	100
	Met Leu Leu Ser Cys Leu Met Phe Leu Ser Trp Val Glu Gly Glu Glu	
	15 20 25	
20	TCT CAA AAG AAA CTG CCT TCT TCA CGT ATA ACC TGT CCT CAA GGC TCT	148
	Ser Gln Lys Lys Leu Pro Ser Ser Arg Ile Thr Cys Pro Gln Gly Ser	
	30 35 40	
25	GTA GCC TAT GGG TCC TAT TGC TAT TCA CTG ATT TTG ATA CCA CAG ACC	196
	Val Ala Tyr Gly Ser Tyr Cys Tyr Ser Leu Ile Leu Ile Pro Gln Thr	
	45 50 55	
30	TGG TCT AAT GCA GAA CTA TCC TGC CAG ATG CAT TTC TCA GGA CAC CTG	244
	Trp Ser Asn Ala Glu Leu Ser Cys Gln Met His Phe Ser Gly His Leu	
	60 65 70 75	
35	GCA TTT CTT CTC AGT ACT GGT GAA ATT ACC TTC GTG TCC TCC CTT GTG	292
	Ala Phe Leu Leu Ser Thr Gly Glu Ile Thr Phe Val Ser Ser Leu Val	
	80 85 90	
40	AAG AAC AGT TTG ACG GCC TAC CAG TAC ATC TGG ATT GGA CTC CAT GAT	340
	Lys Asn Ser Leu Thr Ala Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp	
	95 100 105	
45	CCC TCA CAT GGT ACA CTA CCC AAC GGA AGT GGA TGG AAG TGG AGC AGT	388
	Pro Ser His Gly Thr Leu Pro Asn Gly Ser Gly Trp Lys Trp Ser Ser	
	110 115 120	
50	TCC AAT GTG CTG ACC TTC TAT AAC TGG GAG AGG AAC CCC TCT ATT GCT	436
	Ser Asn Val Leu Thr Phe Tyr Asn Trp Glu Arg Asn Pro Ser Ile Ala	
	125 130 135	
55	GCT GAC CGT GGT TAT TGT GCA GTT TTG TCT CAG AAA TCA GGT TTT CAG	484
	Ala Asp Arg Gly Tyr Cys Ala Val Leu Ser Gln Lys Ser Gly Phe Gln	
	140 145 150 155	
60	AAG TGG AGA GAT TTT AAT TGT GAA AAT GAG CTT CCC TAT ATC TGC AAA	532
	Lys Trp Arg Asp Phe Asn Cys Glu Asn Glu Leu Pro Tyr Ile Cys Lys	
	160 165 170	
65	TTC AAG GTC TAGGGCAGTT CTAATTTCAA CAGCTTGAAA ATATTATGAA	581
	Phe Lys Val	
70	GCTCACATGG ACAAGGAAGC AAGTATGAGG ATTCACCTCAG GAAGAGCAAG CTCTGCCTAC	641
75	ACACCCACAC CAATTCCCTT ATATCATCTC TGCTGTTTTT CTATCAGTAT ATTCTGTGGT	701

- 26 -

GGCTGTAACC TAAAGGCTCA GAGAACAAAA ATAAATGTC ATCAAC

747

(2) INFORMATION FOR SEQ ID NO:2:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 174 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

15

Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp Met Leu Leu Ser Cys
1 5 10 15Leu Met Phe Leu Ser Trp Val Glu Gly Glu Glu Ser Gln Lys Lys Leu
20 25 30

20

Pro Ser Ser Arg Ile Thr Cys Pro Gln Gly Ser Val Ala Tyr Gly Ser
35 40 45

25

Tyr Cys Tyr Ser Leu Ile Leu Ile Pro Gln Thr Trp Ser Asn Ala Glu
50 55 60Leu Ser Cys Gln Met His Phe Ser Gly His Leu Ala Phe Leu Leu Ser
65 70 75 80

30

Thr Gly Glu Ile Thr Phe Val Ser Ser Leu Val Lys Asn Ser Leu Thr
85 90 95

35

Ala Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp Pro Ser His Gly Thr
100 105 110Leu Pro Asn Gly Ser Gly Trp Lys Trp Ser Ser Ser Asn Val Leu Thr
115 120 125

40

Phe Tyr Asn Trp Glu Arg Asn Pro Ser Ile Ala Ala Asp Arg Gly Tyr
130 135 140Cys Ala Val Leu Ser Gln Lys Ser Gly Phe Gln Lys Trp Arg Asp Phe
145 150 155 160

45

Asn Cys Glu Asn Glu Leu Pro Tyr Ile Cys Lys Phe Lys Val
165 170

(2) INFORMATION FOR SEQ ID NO:3:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 175 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(ii) MOLECULE TYPE: protein

(iv) ANTI-SENSE: NO

60

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Rattus rattus

- 27 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

5 Met Leu His Arg Leu Ala Phe Pro Val Met Ser Trp Met Leu Leu Ser
 1 5 10 15
 Cys Leu Met Leu Leu Ser Gln Val Gln Gly Glu Asp Ser Pro Lys Lys
 20 25 30
 10 Ile Pro Ser Ala Arg Ile Ser Cys Pro Lys Gly Ser Gln Ala Tyr Gly
 35 40 45
 Ser Tyr Cys Tyr Ala Leu Phe Gln Ile Pro Gln Thr Trp Phe Asp Ala
 50 55 60
 15 Glu Leu Ala Cys Gln Lys Arg Pro Glu Gly His Leu Val Ser Val Leu
 65 70 75 80
 Asn Val Ala Glu Ala Ser Phe Leu Ala Ser Met Val Lys Asn Thr Gly
 85 90 95
 20 Asn Ser Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp Pro Thr Leu Gly
 100 105 110
 Gly Glu Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Asn Asp Ile Met
 115 120 125
 Asn Tyr Val Asn Trp Glu Arg Asn Pro Ser Thr Ala Leu Asp Arg Gly
 130 135 140
 30 Phe Cys Gly Ser Leu Ser Arg Ser Ser Gly Phe Leu Arg Trp Arg Asp
 145 150 155 160
 Thr Thr Cys Glu Val Lys Leu Pro Tyr Val Cys Lys Phe Thr Gly
 165 170 175

(2) INFORMATION FOR SEQ ID NO:4:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 175 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

45 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

50 Met Leu Pro Pro Met Ala Leu Pro Ser Val Ser Trp Met Leu Leu Ser
 1 5 10 15
 55 Cys Leu Met Leu Leu Ser Gln Val Gln Gly Glu Glu Pro Gln Arg Glu
 20 25 30
 Leu Pro Ser Ala Arg Ile Arg Cys Pro Lys Gly Ser Lys Ala Tyr Gly
 35 40 45

- 28 -

Ser His Cys Tyr Ala Leu Phe Leu Ser Pro Lys Ser Trp Thr Asp Ala
 50 55 60
 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly Asn Leu Val Ser Val Leu
 5 65 70 75 80
 Ser Gly Ala Glu Gly Ser Phe Val Ser Ser Leu Val Lys Ser Ile Gly
 85 90 95
 10 Asn Ser Tyr Ser Tyr Val Trp Ile Gly Leu His Asp Pro Thr Gln Gly
 100 105 110
 Thr Glu Pro Asn Gly Glu Gly Trp Glu Trp Ser Ser Ser Asp Val Met
 115 120 125
 15 Asn Tyr Phe Ala Trp Glu Arg Asn Pro Ser Thr Ile Ser Ser Pro Gly
 130 135 140
 His Cys Ala Ser Leu Ser Arg Ser Thr Ala Phe Leu Arg Trp Lys Asp
 145 150 155 160
 Tyr Asn Cys Asn Val Arg Leu Pro Tyr Val Cys Lys Phe Thr Asp
 165 170 175

25 (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 174 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 30

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Rattus rattus
 35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

40 Met Leu Pro Arg Val Ala Leu Thr Thr Met Ser Trp Met Leu Leu Ser
 1 5 10 15
 Ser Leu Met Leu Leu Ser Gln Val Gln Gly Glu Asp Ala Lys Glu Asp
 20 25 30
 45 Val Pro Thr Ser Arg Ile Ser Cys Pro Lys Gly Ser Arg Ala Tyr Gly
 35 40 45
 50 Ser Tyr Cys Tyr Ala Leu Phe Ser Val Ser Lys Ser Trp Phe Asp Ala
 50 55 60
 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Val Leu
 65 70 75 80
 55 Ser Gly Ser Glu Ala Ser Phe Val Ser Ser Leu Ile Lys Ser Ser Gly
 85 90 95
 Asn Ser Gly Gln Asn Val Trp Ile Gly Leu His Asp Pro Thr Leu Gly
 100 105 110
 60 Gln Glu Pro Asn Arg Gly Gly Trp Glu Trp Ser Asn Ala Asp Val Met

- 29 -

115 120 125
 Asn Tyr Phe Asn Trp Glu Thr Asn Pro Ser Ser Val Ser Gly Ser His
 130 135 140
 5 Cys Gly Thr Leu Thr Arg Ala Ser Gly Phe Leu Arg Trp Arg Glu Asn
 145 150 155 160
 10 Asn Cys Ile Ser Glu Leu Pro Tyr Val Cys Lys Phe Lys Ala
 165 170

(2) INFORMATION FOR SEQ ID NO:6:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 174 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Rattus rattus

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Leu Pro Arg Leu Ser Phe Asn Asn Val Ser Trp Thr Leu Leu Tyr
 1 5 10 15
 30 Tyr Leu Phe Ile Phe Gln Val Arg Gly Glu Asp Ser Gln Lys Ala Val
 20 25 30
 Pro Ser Thr Arg Thr Ser Cys Pro Met Gly Ser Lys Ala Tyr Arg Ser
 35 35 40 45
 Tyr Cys Tyr Thr Leu Val Thr Thr Leu Lys Ser Trp Phe Gln Ala Asp
 50 50 55 60
 Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Ile Leu Ser
 40 65 70 75 80
 Gly Gly Glu Ala Ser Phe Val Ser Ser Leu Val Thr Gly Arg Val Asn
 85 90 95
 45 Asn Asn Gln Asp Ile Trp Ile Trp Leu His Asp Pro Thr Met Gly Gln
 100 105 110
 Gln Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Ser Asp Val Leu Asn
 50 115 120 125
 Tyr Leu Asn Trp Asp Gly Asp Pro Ser Ser Thr Val Asn Arg Gly Asn
 55 130 135 140
 Cys Gly Ser Leu Thr Ala Thr Ser Glu Phe Leu Lys Trp Gly Asp His
 145 150 155 160
 His Cys Asp Val Glu Leu Pro Phe Val Cys Lys Phe Lys Gln
 165 170

60 (2) INFORMATION FOR SEQ ID NO:7:

- 30 -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 165 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Rattus rattus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Thr Arg Asn Lys Tyr Phe Ile Leu Leu Ser Cys Leu Met Val Leu
1 5 10 15
Ser Pro Ser Gln Gly Gln Glu Ala Glu Glu Asp Leu Pro Ser Ala Arg
20 25 30
Ile Thr Cys Pro Glu Gly Ser Asn Ala Tyr Ser Ser Tyr Cys Tyr Tyr
35 40 45
Phe Met Glu Asp His Leu Ser Trp Ala Glu Ala Asp Leu Phe Cys Gln
50 55 60
Asn Met Asn Ser Gly Tyr Leu Val Ser Val Leu Ser Gln Ala Glu Gly
65 70 75 80
Asn Phe Leu Ala Ser Leu Ile Lys Glu Ser Gly Thr Thr Ala Ala Asn
85 90 95
Val Trp Ile Gly Leu His Asp Pro Lys Asn Asn Arg Arg Trp His Trp
100 105 110
Ser Ser Gly Ser Leu Phe Leu Tyr Lys Ser Trp Asp Thr Gly Tyr Pro
115 120 125
Asn Asn Ser Asn Arg Gly Tyr Cys Val Ser Val Thr Ser Asn Ser Gly
130 135 140
Tyr Lys Lys Trp Arg Asp Asn Ser Cys Asp Ala Gln Leu Ser Phe Val
145 150 155 160
Cys Lys Phe Lys Ala
165